QUADERNI

Università degli Studi di Siena
DIPARTIMENTO DI ECONOMIA POLITICA

ERNESTO SAVAGLIO

Multivariate Differences Ordering

n. 419 - Gennaio 2004

Abstract: A multivariate dispersion ordering is introduced as a natural extension of a well-known univariate inequality criterion. A characterization of this new ordering is provided.

JEL Classification: I31

Keywords: Inequality, Orderings, Multidimensional Distributions.

Acknowledgements: I thank A. Battinelli and an anonymous referee for comments and suggestions. The usual disclaimer applies.

Ernesto Savaglio, Dipartimento di Economia Politica, Università di Siena

1. Introduction

The economic literature concerning inequality measurement has recently extended the analysis from a unidimensional setting, where income (or wealth) only matters, to a multidimensional framework, where many individual characteristics have taken into account. In fact, in order to evaluate the social state of a population of individuals, who differ in many aspects besides income, more than one criterion needs to be applied. In this note, we introduce the multivariate generalization of a well-known univariate inequality criterion that takes into account all changes that could occur in every quantile of a (income) distribution. In the following subsections we first introduce our setting, therefore we compare the unidimensional ordering that we go to generalise with the Lorenz criterion, the most utilized criterion for studying income inequality. In section 2, we define the multivariate differences ordering (simply MD). Finally, we provide a simple characterization of MD, we relate MD to a well-known ordering using doubly stochastic matrices and conclude.

1.1. **Notation and Definitions.** Let us consider a random variable $\mathbf{z}$ that can be interpreted as an income distribution, and define the cumulative distribution function of $\mathbf{z}$ as follows:

$$F(z) = \sum_{\{i|z_i \leq z\}} p_i \quad \forall z \in \mathbb{R},$$

where $F(z)$ could be interpreted as the percentage of people in the distribution $\mathbf{z}$ receiving income less than or equal to the income z. The left continuous version of the inverse of F, the so-called quantile function, is a function $F^{-1} : [0,1] \rightarrow \mathbb{R}_+$, that is denoted as:

$$F^{-1}(p) = Inf\left\{z \in \mathbb{R}_+, F(z) \geq p : p \in [0,1]\right\}.$$

The inequality criterion considered in this work is defined as follows:[1]

Definition 1. *Given two (income) distributions* $\mathbf{x}$*,* $\mathbf{y}$*, we say that* $\mathbf{y}$ *is less unequal than* $\mathbf{x}$ *for the differences ordering, denoted as* $\mathbf{y} \preceq_D \mathbf{x}$*, if*

$$F_{\mathbf{y}}^{-1}(v) - F_{\mathbf{y}}^{-1}(u) \leq F_{\mathbf{x}}^{-1}(v) - F_{\mathbf{x}}^{-1}(u), \ \forall\ 0 < u < v \leq 1. \tag{1.1}$$

Fraser [3] first introduced Definition 1. Karlin [4] analyses 1.1 in a classical work on stochastic orderings. Marshall et al. [6] discusses such a ordering in a setting of Theory of Majorization. In an unpublished work, Preston [7] introduces Definition 1 in economics, stressing the possibility to replace the Lorenz ordering (henceforth L), with this alternative criterion.

1.2. **On the relation between D and L.** Let us first define the Lorenz criterion as follows:

Definition 2. *A distribution* $\mathbf{y}$ *is less unequal than* $\mathbf{x}$ *according to Lorenz ordering, denoted* $\mathbf{y} \preceq_L \mathbf{x}$*, if and only if*

$$F_{\mathbf{y}}^{-1}(v) \geq F_{\mathbf{x}}^{-1}(v) \textit{ for any } v \in [0,1).$$

Several scholars have noticed that, in some situations, L fails as a suitable inequality criterion (see Savaglio [8] and references therein). They therefore introduced D as a suitable alternative. The following proposition analyses the relation

[1] An *inequality criterion* $\preceq$ is a *partial ordering*, i.e. an asymmetric and transitive binary relation. When two income distributions $\mathbf{x}$, $\mathbf{y}$ satisfy $\mathbf{y} \preceq \mathbf{x}$, we shall say that $\mathbf{y}$ is less unequal than $\mathbf{x}$.

between D and L by relating these two inequality criteria to the *sign-changed* orderings.[2]

Let F and G be the cumulative distributions respectively of $\mathbf{x}$, $\mathbf{y}$. A common way to compare the dispersion (inequality) of F and G is via the sign changes of $(F - G)$. If $S(h)$ is the number of sign changes of the function $h(t)$, a natural condition on $(F - G)$ corresponding to F being in some sense more variable than G is:

$$S(F - G) = 1 \quad \text{with sign sequence } +\,,\ -\,.$$

Further, let $\varphi(y) = F^{-1}(G(y))$ be a continuous non-decreasing function defined on $[0, 1]$ and onto, then:

Proposition 1. *If* $\mathbf{y} \preceq_D \mathbf{x}$ *then* $\mathbf{y} \preceq_L \mathbf{x}$.

Proof. It is known that $\mathbf{y} \preceq_L \mathbf{x}$ if and only if $S(F - G) = 1$ (see Karlin [4] vol. I, chapter 5), and that $\mathbf{y} \preceq_D \mathbf{x}$ if and only if $S(F - G) \leq 1$ (see Shaked [9] Theorem 2.1). This means that D implies L while the contrary does not necessary hold.

According to Marshall and Olkin [5], $\mathbf{y} \preceq_D \mathbf{x}$ is tantamount to $F^{-1}(G(y))$ non-decreasing on $[0, 1]$. A necessary and sufficient condition for $S(F - G) \leq 1$ is that the two distribution functions F and G cut one another at most once. Now, $F^{-1}(G(y))$ crosses any line $\mathbf{x}(x) = kx$ *at most once*, and from below, and, having crossed it, never touches it again. Taking $k = 1$ and using the non-decreasing nature of F^{-1}, it follows that, if there exists a crossing point ξ, then:

$$G(y_0) \gtreqless F(y_0) \quad \text{as} \quad y_0 \lesseqgtr \xi.$$

[2] See Karlin (1968) vol. I, chapters 5 and 6.

This means that F and G cut one another at most once and then $S(F-G) \leq 1$ as required.[3] □

As D implies L, it means that 1.1 compares less distributions' pairs than Lorenz ordering does and that each pair compared by D is compared by L. Hence, the conclusion that we get is not so encouraging: D could not be a suitable alternative to L, but at most a complementary ordering checking the information about the distribution's quantiles. Nevertheless, as D is a suitable inequality criterion (see Savaglio [8] for the analysis of some interesting properties of D), extending such a ordering to a multidimensional context is worth pursuing. Next section is then devoted to the study of the differences ordering when individuals differ in many aspects besides income.

2. Multivariate differences ordering

Economic disparity does not arise from the distribution of income alone. People are different in income, education, health, etc. and we must take into account several individual characteristics if we want to understand and evaluate inequality among people. For such a reason, we extend our measurement to several variables, in order to consider the other attributes (e.g. health, education, talents, capabilities etc.), that characterize individuals. In order to generalise D, we consider multivariate distributions representing populations of individuals with different characteristics whose distributions are random variables. Let us introduce first the following:

[3] Alternative proofs of this well-known result are provided by e.g. Marshall and Olkin [5] proposition B.1 on page 129 and Savaglio [8].

Definition 3. *A function $f : \mathcal{A} \subseteq \mathbb{R}^n \to \mathbb{R}^n$, which satisfies the Lipschitz condition $\left\| f(\mathbf{x}) - f\left(\mathbf{x}'\right) \right\| \leq \left\| \mathbf{x} - \mathbf{x}' \right\|$, with $\left(\mathbf{x}, \mathbf{x}'\right) \in \mathcal{A} \times \mathcal{A}$ and $\|\cdot\|$ the Euclidean norm, is a contraction mapping.*

A multidimensional distribution X (or, equivalently, a random vector), is a list of d (column) random variables $X = \left(\mathbf{x}^1, ..., \mathbf{x}^d\right)$. We interpret $\mathbf{x}^j$ $(j = 1, ..., d)$, as the distribution of the jth characteristics among n individuals. For any two random vectors X and Y, we suppose that $\mathbf{x}^j \in X$ has the same average as $\mathbf{y}^j \in Y$ for all $j = 1, ..., d$. By converse, $\mathbf{x}_i$ $(i = 1, ..., n)$, is the (row) random variable representing the ith individual distribution of d characteristcs. Denote with $\aleph$ the set of all vectors of random variables with the same characteristics' average. Hence, we define the *multivariate difference ordering* (MD) as follows:

Definition 4 (MD)**.** *Let X and Y be two multidimensional distributions in $\aleph$. Then Y is said to be multidimensional differences majorized by X, denoted as $Y \preceq_{MD} X$, if and only if there exists a contraction function $k(\cdot)$ such that Y has the same distribution as $k(X)$, namely $Y \sim k(X)$ and*

$$\|k(\mathbf{x}_{i+1}) - k(\mathbf{x}_i)\| \leq \|\mathbf{x}_{i+1} - \mathbf{x}_i\| \tag{2.1}$$

for all $\mathbf{x}_{i+1}$, $\mathbf{x}_i \in X$.

In other words, a contraction function is an inequality reducing transformation that makes a distribution smoother. Suppose to apply a progressive tax characteristic by characteristic and equally redistribute the taxation's amount received still characteristic by characteristic. Such a transformation is mean-preserving and inequality-reducing. When a redistributive policy like this occurs the individual d-characteristics distributions will be smoother for each individual i, with $i = 1, ..., n$.

The final result of such an equality-enhancing process is analytically captured by the contraction function $k(\cdot)$.

There has been a great deal of interest in studying maps preserving a given preordering $\preceq$ on a set $\wp$, i.e. real-valued functions φ satisfying:

$$\varphi(\mathbf{Y}) \leq \varphi(\mathbf{X}) \qquad \text{whenever } \ \mathbf{Y} \preceq \mathbf{X} \text{ with } X, Y \in \wp.$$

Functions φ are variously referred to as monotonic, isotonic or order-preserving. In what follows, we look for *linear* maps that preserve the ordering of MD. In order to provide such a characterization, few basic definitions and results are needed.

Definition 5. *Let $\wp$ be a set of all $n \times n$ positive semidefinite matrices in $\mathbb{R}^{n \times n}$. Then Y is said to be Lowner majorized by X, denoted as $Y \preceq_{Low} X$, if and only if $(X - Y)$ is non-negative definite, i.e. belongs to the convex cone of positive semindefinite matrices in $\wp$ for any $X, Y \in \wp$.*

Denote with I_n the n identity matrix and $\mathbb{J}_k(\mathbf{z}) = \{\partial k_i / \partial z_j\}$ the Jacobian matrix of a $k : \mathbb{R}^n \to \mathbb{R}^n$ contraction function. According to Marshall and Olkin [5], a continuous differentiable function $k(\cdot)$ is a contraction if and only if $\mathbb{J}_k(\mathbf{z})^T \mathbb{J}_k(\mathbf{z}) \preceq_{Low} I_n$ for all $\mathbf{z} \in \mathbb{R}^n$. Hence, it is obvious that if Y has the same distribution as $k(X)$, with $k(\cdot)$ continuously differentiable, then $\mathbb{J}_k(\mathbf{z})^T \mathbb{J}_k(\mathbf{z}) \preceq_{Low} I_n$ if and only if $Y \preceq_{MD} X$. Now, consider the following result of Eaton [2] shows that

Proposition 2. *For any square matrix H, $H^T H \preceq_{Low} I_n$ if and only if $H = \sum_{i=1}^{m} \alpha_i \Gamma_i$ for some orthogonal matrices Γ_i, all $\alpha_i > 0$ and $\sum_{i=1}^{m} \alpha_i = 1$.*

Then, we immediately state the following characterization:

Theorem 1. *If Y has the same distribution as $AX+d$ (i.e. $Y \sim AX+dm$), for a fixed matrix A and a fixed n-vector d, then $Y \preceq_{MD} X$ if and only if $A=\sum_{i=1}^{m}\alpha_i\Gamma_i$ with $\alpha_i>0$ and $\sum_{i=1}^{m}\alpha_i=1$, and Γ_i are orthogonal matrices.*

A special case of representation of the MD-ordering is when Y has the same distribution as DX with D a doubly stochastic matrix.[4]. Finally, we show a sufficient condition for a linear invertible map to preserve MD.

Theorem 2. *Suppose A be a n-matrix such that for any orthogonal matrix Γ there is an orthogonal matrix Γ^{-} such that $\Gamma A\Gamma^{-}=A$. Then $Y \preceq_{MD} X$ implies $AY \preceq_{MD} AX$.*

Proof. Assume Y have the same distribution as $k(X)$ where $k(\cdot)$ is a contraction. Then, $\mathbb{J}_k(\mathbf{x})=\sum_{i=1}^{n}\alpha_i\Gamma_i(x)$ and $A\mathbb{J}_kA^{-1}=\sum_{i=1}^{n}\alpha_iA\Gamma_iA^{-1}=\sum_{i=1}^{n}\alpha_i\Gamma_i^{-}AA^{-1}$, and hence AkA^{-1} is a contraction. □

When A reduces to an n-vector $\mathbf{a}$, we get what is called by Bandhari [1] *directional majorization* and that Kolm calls *price majorization.* Theorem 2 can therefore be interpreted in terms of prices and expenditures, saying that Y has less multivariate inequality than X if and only if $\mathbf{a}Y$ has less univariate inequality than $\mathbf{a}X$ for every $\mathbf{a}\in\mathbb{R}_{+}^{d}$, where d is the dimension of the individual characteristics.

2.1. **Conclusion.** The problem to study inequality in a context of more than one variable is inherently complex. The principal reason of such a difficulty is relative to the interaction between income and non-income attributes. In the present work, we have not considered such a problem of correlation. We have extended a disparity

[4] A square matrix A is said to be *doubly stochastic* if its elements are all non-negative and all row and column sums are one.

criterion to a multivariate context in order to evaluate the social state of a population of individual who differ in many characteristics. Further, we have provided a simple characterization of this multidimensional inequality criterion. Finally, the relation between such a new ordering and the directional majorization represents an insight for also comparing qualitative variables in such unexplored field.

References

[1] Bandhari, S.K. (1988) - Multivariate majorization and directional majorization: positive results - *Sankhyā*, Ser. A, 50, 199-204.

[2] Eaton, M.L. (1984) - On group induced orderings, monotone functions and convolutions theorems - in : Y.L. Tong ed., *Inequalities in statistics*, IMS lectures Notes.

[3] Fraser, D.A.S. (1957) - *Nonparametric Methods in Statistics* - Wiley, New York.

[4] Karlin, S. (1968) - *Total Positivity* - vol. I

[5] Marshall, A.W., and I. Olkin, (1979) - *Inequalities: Theory of Majorization and its Applications* - Academic Press: New York.

[6] Marshall, A.W., I. Olkin, and F. Proschan, (1967) - Monotonicity of ratio of means and other applications of majorization - in *Inequalities* (O. Shisha, ed.) pp.177-190. Academic Press, New York.

[7] Preston, I. (1990) - Ratios, differences, and inequality indices - Institute for Fiscal Studies, Working Paper W90-9.

[8] Savaglio, E. (2003) - A Note on Inequality Criteria - *Mathematical Inequalities and Applications*, vol.6.

[9] Shaked, M. (1982) - Dispersive ordering of distributions - Journal of Applied Probability 10, 310-320.